Machine Learning

Stanford University
Professor Andrew Ng

Jordan Hong

May 22, 2020

Contents

1 Introduction **1**
1.1 What is Machine Learning 1
1.2 Supervised Learning 2
1.3 Unsupervised Learning 2

2 Linear Regression with One Variable **3**
2.1 Model Representation 3
2.1.1 Notations 3
2.1.2 Hypothesis Function 3
2.2 Cost Function 3
2.3 Gradient Descent 4

1 Introduction

1.1 What is Machine Learning

1. Machine Learning
   - Grew out of work in Artificial Intelligence (AI)
   - New capabilities for computers
2. Examples:
   - database mining
   - applications can't programby hand (handwriting recognition, Natural Language Processing (NLP), Computer Vision)
   - Neuromorphic applications
3. Definition

- Arthur Samuel(1959)

  Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.

- Tom Mitchell(1998)

  Well-posed Learning Problem: A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

4. Machine Learning in this course:
   (a) Suupervised Learning
   (b) Unsupervised Learning
   (c) Others: reinforcement learning, recommender systems
   (d) Practical application techniques

1.2 Supervised Learning

In supervised learning, the *the right answer* is given. For example:

1. Regression: predict real-valued output.
2. Classification: predict discrete-valued output.

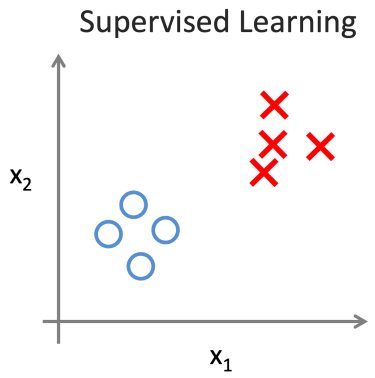

Figure 1: Supervised Learning

1.3 Unsupervised Learning

The right answer is not given, e.g. cocktail problem (distinguishing two voices from an audio file.)

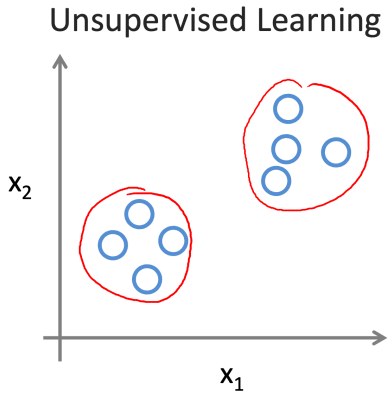

Figure 2: Unsupervised learning

2 Linear Regression with One Variable

2.1 Model Representation

2.1.1 Notations

For a training set:

- **m** = Number of training examples.
- **x** = "input" variable / features.
- **y** = "output" variables / "target" variable.
- **(x,y)** - one training example.
- $\mathbf{(x^i,y^i)}$ denotes the i^{th} training example

2.1.2 Hypothesis Function

A hypothesis function (h) maps input (x) to estimated output (y). How do we represent h?

$$h_\theta(x) = \theta_0 + \theta_1 \times x \tag{1}$$

We can apply *Univariate linear regression* with respect to x.

2.2 Cost Function

Recall 1. The θ^is are parameters we have to choose. The intuition is is that we want to choose θ_i s such that h_θ is closest to y for our training examples (x,y).
Cost Function:

2.3 Gradient Descent